The design of experiments

Statistical principles for practical application

R. MEAD

The design of experiments

STATISTICAL PRINCIPLES FOR PRACTICAL APPLICATIONS

R. MEAD
Department of Applied Statistics, University of Reading

CAMBRIDGE UNIVERSITY PRESS
Cambridge
New York New Rochelle Melbourne Sydney

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
32 East 57th Street, New York, NY 10022, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1988

Printed in Great Britain at the University Press, Cambridge

British Library cataloguing in publication data

Mead, R.
The design of experiments: statistical
principles for practical applications.
1. Experimental design
I. Title
001.4′34′028 QA279

Library of Congress cataloguing in publication data

Mead, R. (Roger)
The design of experiments.

Bibliography
Includes index.
1. Experimental design. 2. Mathematical statistics.
I. Title.
QA279.M388 1988 001.4′34 87-24236

ISBN 0 521 24512 5 hard covers

DE

PREFACE

My aim in this book is to explain the fundamental statistical concepts required for designing efficient experiments to answer real questions. The book has grown out of 25 years experience of designing experiments for research scientists, and of teaching the concepts of statistical design both to statisticians and to experimenters. This experience has convinced me that the whole subject of statistical design needs to be reassessed in the context of modern statistical computing facilities. The development of statistical philosophy about the design of experiments has been dominated during the last 30 years by mathematical theory. The influence of the availability of vastly improved computing facilities on teaching, textbooks, and, most crucially, practical experimentation appears to have been slight.

The existence of statistical programs capable of analysing the results from any experimental design does not imply any changes in the main statistical concepts of design. However these concepts have become restricted by the earlier need to develop mathematical theory for design in such a way that the results from the designs can be analysed without recourse to computers. The fundamental concepts now require re-examination and re-interpretation outside the limits implied by classical mathematical theory so that the full range of design possibilities may be considered. The result of the revolution in computing facilities is that experimental design should become a much wider and more exciting subject. I hope that this book will display that breadth and excitement.

The development of this book has been particularly motivated by teaching postgraduate students specialising in statistics. However the intention of the book encompasses a wider audience. Understanding the fundamental concepts of design is essential for all research scientists involved in programmes of experimental work. In addition to this general need for an understanding of the philosophy of experimental design there are particular aspects of design, such as repeated measurements (Chapter 14) or levels of replication (Chapter 6) which are relevant to virtually all research disciplines in which experimental work is required. Because of the concentration on basic concepts and their implications the book could be used for courses for final-year undergraduates provided such courses allow sufficient time for the concepts to be thoroughly discussed.

I have tried in writing this book to concentrate on the ideas of design rather than those of analysis, on the statistical concepts rather than the mathematical theory, and on practically useful experiments rather than classes of possible experimental designs. Obviously it is necessary to know how to analyse data from experiments and the philosophy and methods of analysis are discussed in the introductory first part of the book. Also of course, examples of analysis punctuate many later sections. However, after the first part of the book it is assumed that the analysis of experimental data is not difficult when modern computing facilities are used. Consequently ideas of analysis are introduced only when they illumine or motivate the design concepts.

The formal language of statistics is mathematical. It is not possible to discuss design without some mathematically complex formulation of models and ideas. Some of the mathematical language used in the book requires a sound mathematical background beyond school level. However in all parts of the book it is the statistical concepts which are important, and the structure of the book hopefully allows the less-mathematical reader to bypass the more complex mathematical details. Throughout the book the development of concepts relies on many examples. I hope that readers will consider the detailed arguments of these examples. By trying to solve the problems which underly the examples before reading through the explanation of the solutions, I believe that readers will start to develop the intuitive understanding of design concepts which is essential to good design. For the mathematically sophisticated reader the mathematical details provide additional support for the statistical concepts.

Most importantly, the book is intended to show how practical problems of designing real experiments should be solved. To stimulate this practical emphasis real examples of design problems are described at the beginning of each chapter from Chapter 6 onwards (except Chapter 9). The final chapter of the book attempts an overall view of the problem-solving aspects of design. The bias of the areas of application discussed in this book inevitably reflects my personal experience. Much of my experience has been concerned with agricultural experimentation and to write with anything like a comprehensive knowledge about the practical application of design concepts I believe that a concentration on agricultural applications is inevitable. However all the examples are intended to illustrate particular forms of problem that will be relevant in many other fields of application. I hope that statisticians and research scientists in a wide range of experimental disciplies will be able to interpret and adapt the concepts discussed in the book to their own requirements through the use of analogy when the examples discussed are not directly relevant to their discipline.

The book is divided into an overture and two main subjects with a final coda to bring together all the previous material. Chapters 1 to 5 constitute the overture, providing a general introduction and the basic theory necessary for analysis of experimental data. Chapters 1, 2 and 3 should be familiar to readers who have taken an elementary course in experimental design. Alternatively, for those readers without any previous training in design, these three chapters provide an introductory presentation of the two most important ideas of design, blocking and factorial structure. Chapter 4 is the mathematically heavy chapter, providing the necessary theory for general linear models and the analysis of data from designed experiments, with an initial explanation of the important results at a rather simpler level. Chapter 5 ventures briefly, and rashly, into an inevitably prejudiced view of computing needs and implications.

The first main subject is unit variation and control. The fundamental concepts of replication, blocking (with either one or two systems of control) and randomisation are each examined separately. My aim is to distinguish the purposes and practical relevance of each concept and to eliminate the confusion about these concepts which seems to be common in the minds of users of experimental designs. The other two chapters in this subject, on covariance (for control) and on assumptions (to express variation realistically) are biased towards analysis rather than design but their implications for design and for the choice of measurement variables are important.

The second main subject is treatment questions and structure. Chapter 12 presents a broad view of the need for statisticians to be involved in all stages of discussion about the choice of treatments and the interpretation of results. The classical ideas of factorial structure and single and fractional replicates are presented in Chapter 13. Important consequences of particular practical requirements for factorial structure are described in Chapters 14 (split levels of information) and 15 (avoiding confounding). Some necessary mathematical theory for confounding is included in Chapter 16. The choice of experimental treatments for the investigation of the response to quantitative factors is discussed in Chapters 17 and 18.

Finally in the coda, Chapter 0 seeks to draw the concepts of the two main subjects together to provide guidance on choosing experimental designs to satisfy particular practical requirements. A book on practical experimental design should start with the approach of Chapter 0 but this chapter requires knowledge from the previous chapters before it can be read and understood. Hence the number and position of this chapter.

I owe a considerable debt to many consultees and collaborators both for the stimulus to consider why the problems they presented should be covered by a book on design and also for the many examples they have provided. There are too many for me to thank them individually here for their stimulating requests and they therefore remain anonymous (some should prefer it that way, and others are too far into the mists of time for anything else). I have also drawn exercises at the end of the chapters from Chapter 6 onwards from many examination papers.

I have also benefitted from many discussions with colleagues at Reading and wish to thank Richard Coe, Robert Curnow, Derek Pike and Roger Stern, without whom this book would have been a more stunted growth. I am particularly grateful to Richard Jarrett for a most stimulating dinner conversation in Melborne, from which Chapter 0 was born.

Finally I record my gratitude to those who have made the production of this book possible. Clive Bowman and Marilyn Allum manipulated GLIM with great skill to provide many analyses and information about precision for possible designs. Audrey Wakefield has worked devotedly through thousands of pages of manuscript, second thoughts, third thoughts, total reorganisations . . . , and has endured the slings and arrows of an early generation word-processor. Rosemary Stern has used her art to convert ideas into figures. And David Tranah and Martin Gilchrist at Cambridge University Press have been most patient with a tardy author.

CONTENTS

PART I

OVERTURE

1

Introduction

1.1 Why a statistical theory of design?

The need to develop statistical theory for designing experiments stems, like the need for statistical analysis of numerical information, from the inherent variability of experimental results. In the physical sciences, this variability is frequently small and, when thinking of experiments at school in physics and chemistry, it is usual to think of 'the correct result' from an experiment. However, practical experience of such experiments makes it obvious that the results are, to a limited extent, variable, this variation arising as much from the complexities of the measurement procedure as from the inherent variability of experimental material. As the complexity of the experiment increases, and the differences of interest become relatively smaller, then the precision of the experiment becomes more important. An important area of experimentation within the physical sciences, where precision of results and hence the statistical design of experiments are important, is the control of industrial chemical processes.

Whereas the physical sciences are thought of as exact, it is quite obvious that biological sciences are not. Most experiments on plants or animals use many plants or animals because it is clear that the variation between two plants, or between two animals, is very large. It is impossible, for example, to predict quantitatively the characteristics of one plant from the corresponding characteristics of another plant of the same species, age and origin.

Thus, no medical research worker would make confident claims for the efficacy of a new drug merely because a single patient responded well to the drug. In the field of market research, no newspaper would publish an opinion poll based on interviews with only two people, but would require a sample of at least 500, together with information about the method of selection of the sample. In a drug trial, the sample of patients would normally be less than 500, possibly between 20 and 100. In psychological experiments, the number of subjects used might be only eight to 12. In agricultural experiments, there may be 20 to 100 plots of land, each with a crop grown on it. In a laboratory experiment hundreds of plants may be treated and examined individually. Or just six cows may be examined while undergoing various diets, with measurements taken frequently and in great detail.

The size of an experiment will vary according to the type of experimental method and the objective of the experiment. One of the important statistical ideas of experimental design is the choice of the size of an experiment. Another is the control of the use of experimental material. It is of little value to use large numbers of patients in the comparison of two drugs, if all the patients given one drug are male, aged between 20 and 30, and all the patients given the other drug are female, aged 50 to 65. Any reasonably sceptical person would doubt claims made about the relative merits of the two drugs from such a trial. This example may seem trivially obvious, but the scientific literature in medicine and many other disciplines shows that many examples of badly planned (or unplanned) experiments occur.

And this is just the beginning of statistical design theory. From avoiding foolish experiment, we can go on to plan improvements in precision for experiments. We can consider the choice of experiments as part of research strategy and can, for example, discuss the relative merits of many small experiments or a few large experiments. We can consider how to design experiments when our experimental material is generally heterogeneous, but includes groups of similar experimental units. Thus, if we are considering the effects of applying different chemicals on the properties of different geological materials, then these may be influenced by the environment from which they are taken, as well as by the chemical treatment applied. However, we may have only two or three samples from some environments, but as many as ten samples from other environments; how then do we decide which chemicals to apply to different samples so that we can compare six different chemical treatments?

1.2 History, computers and mathematics

If we consider the history of experimental design, then most of the developments have been in biological disciplines, in particular in agriculture, and also in medicine and psychology. There is therefore an inevitable agricultural bias to any discussion of experimental design. Most of the important principles of experimental design were developed in the 1920s and 1930s by R. A. Fisher. The practical manifestation of these principles was very much influenced by the calculating capacity then available. Had the computational facilities which we now enjoy been available when the main theory of experimental design was being developed then, I believe, the whole subject of design would have developed very differently. Whether or not this belief is valid, it is certainly true that a view of experimental design today must differ from that of 50, or even 30, years ago. The principles have not changed, but the principles are often forgotten, and only the practical manifestation of the principles retained; these practical applications do require rethinking.

The influence of the computer is one stimulus to reassessing experimental design. Another cause for concern in the development of experimental design is the tendency for increasingly formal mathematical ideas to supplant the statistical ideas. Thus the fact that a particularly elegant piece of mathematics can be used to demonstrate the existence of groups of designs, allocating treatments to blocks of units in a particular way, begs the statistical question of whether such designs would ever be practically useful.

Although myself originally a mathematician, I believe that the presentation of statistical design theory has been quite unnecessarily mathematical, and I shall hope to demonstrate the important ideas of statistical design without excessive mathematical encumbrance. The language of statistical theory, like that of physics, is mathematical and there will be sections of the book where those with a mathematical education beyond school level will find a use for their mathematical expertise. However, even in these sections, which I believe should be included because they will improve the understanding of statistical theory of the readers able to appreciate the mathematical demonstrations, there are intuitive explanations of the theory at a less advanced mathematical level.

1.3 The influence of analysis on design

To write a book solely about the theory of experimental design, excluding all mention of the analysis of data, would be impossible. Any experimenter must know how he intends to analyse his experimental data before he designs his experiment to yield the data. If not, how can he know whether the form of information which he collects can be used to answer the questions which prompted him to do an experiment?

Thus, consider again the medical trial to compare two drugs. Suppose the experimenter failed to think about the analysis and argued that one of the drugs was well known, while the other was not; in the controlled experiment to compare them, there are available 40 patients. Since a lot is